



كلية التربية

المجلـــة العلمـــية

======

الممارسات الإحصائية الخاطئة في حساب صدق وثبات الأدوات في الرسائل العلمية

إعـــــداد

أ.د / عماد أحمد حسن علي

أستاذ علم النفس التربوي

ووكيل كلية التربية – جامعة أسيوط.

أ.د / محمد رياض أحمد عبد الحليم

أستاذ ورئيس قسم علم النفس التربوي

كلية التربية– جامعة أسيوط.

م م/ علي صلاح عبد المحسن حسن

مدرس مساعد بقسم علم النفس

التربوي، كلية التربية– جامعة أسيوط.

﴿ المجلد التاسع والعشرون – العدد الثالث – يوليو ٢٠١٣م ﴾



أ.د / عماد أحمد حسن علي
أ.د / محمد رياض أحمد عبد الحليم
م م/ علي صلاح عبد المحسن حسن

مقدمة البحث:

الإحصاء أسلوب ضروري ومهم لإدخال الدقة والموضوعية في البحوث بكافة أنواعها المختلفة ومجالاتها المتنوعة، وأية بحوث أو دراسات يقوم بها الباحثون يجب أن يستخدموا أساليب إحصائية تتناسب وطبيعة دراستهم، فكثير من البحوث والدراسات يستخدم فيها الباحثون أساليب إحصائية بشكل غير مناسب لطبيعة بيانات هذه البحوث والدراسات، فنجد بعض الباحثين يستخدم إحصاء بارامترياً عند معالجة فروض دراستهم أو حساب كفاءة أدوات هذه الدراسات وكان من الأفضل ـ طبقاً لمحكات ومعايير إحصائية ـ استخدام الإحصاء اللا بارامتري، ومع زيادة أعداد وأنواع البحوث التربوية التي تمت إجازتها من كليات التربية المختلفة تزداد الحاجة إلى ضرورة التعرف على واقع استخدام الأساليب الإحصائية المتنوعة في البحوث والدراسات التربوية، وأهم الأخطاء الشائعة في استخدام الأساليب الإحصائية المختلفة في هذه البحوث والدراسات والعمل على تقديم بعض الإرشادات والتوصيات التي عند الأخذ بها قد يقلل من ظهور مثل هذه الأخطاء في البحوث المقبلة.

ومن الملاحظ أن عملية جمع البيانات في حد ذاتها ليست غاية يسعى إليها الباحث، بل هي وسيلة إلى تحقيق أهداف معينة، أي أن عمل الباحث لا ينتهي عند جمع البيانات، بل ينطلق منها؛ لذلك فالهدف النهائي من اختيار الوسائل المناسبة لجمع البيانات هو الحصول على البيانات التي تخدم البحث وتساعد في دراسة المشكلة وإيجاد الحلول المناسبة لها، وقد تكون جميع مراحل جمع البيانات جيدة وصحيحة، ولا يتخللها ثغرات يمكن أن تؤثر في صحة البيانات، ولكن هذه البيانات تكون مادة خاماً لا تعطي ثمارها ما لم يوجد الباحث المتميز الذي يستطيع استثمارها على الوجه الأنسب، فقد تقع هذه البيانات في أيدي باحثين اثنين أو أكثر، ولكن تحليلهم وتفسيرهم لها يختلف وقد يصل إلى درجة التضاد، وبذلك تختلف النتائج التي يتوصلون إليها وما بني عليها من استنتاجات، وما تسفر عنه من توصيات ومقترحات، لما هناك من فروق بين الباحثين في مدى دقة استخدام واختيار الأسلوب الإحصائي الأكثر ملاءمة مع تلك البيانات المتوافرة.

إن المستقرئ للبحوث والدراسات يجد أن تحليل البيانات وتفسيرها يتأثر بعدة عوامل تتعلق بمدى معرفة الباحثين بالأساليب الإحصائية الملائمة لتحليل بيانات أبحاثهم، ودرجة فهم الباحثين للإحصاءات كرموز لظواهر تكمن وراءها، لذلك فإن الباحثين ما لم يكونوا على دراية بأساليب الإحصاء الملائمة، وفهم عميق لطبيعة البيانات المتعلقة بدراستهم، فإن ذلك سيؤدي إلى نتائج مضللة وغير مفيدة؛ ولإيضاح المشكلات التي تقابل الباحثين تم تصميم هذه البحث لهدف التعرف على طبيعتها ومحاولة إيجاد مخرج للحد من تأثيرها على نتائج الأبحاث.

مشكلة البحث:

تبلورت مشكلة البحث في السؤال الرئيس " ما مدى مناسبة الأساليب الإحصائية المستخدمة في حساب كفاءة الأدوات في بعض الرسائل العلمية بالأقسام التربوية بكلية التربية خلال الفترة من ٢٠٠٤ وحتى ٢٠٠٩ وذلك في ضوء مستويات القياس؟ ". ويتفرع عن هذا السؤال الرئيس عدة أسئلة فرعية هي:

١- ما الأساليب الإحصائية الأكثر شيوعاً في الاستخدام من قِبل الباحثين بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط؟

٢- ما مدى مناسبة الأساليب الإحصائية المستخدمة في حساب كفاءة الأدوات البحثية في الرسائل العلمية بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس؟

أهداف البحث:

يهدف البحث الحالي إلى التعرف على الممارسات الإحصائية الخاطئة في حساب الخصائص السيكومترية للأدوات المستخدمة في رسائل الماجستير والدكتوراه بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس، من خلال:

١- التعرف على الأساليب الإحصائية الأكثر شيوعاً في الاستخدام من قِبل الباحثين بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط.

٢- تقييم الأساليب الإحصائية الأكثر استخداماً من قِبل الباحثين في الحكم على أدوات الرسائل العلمية بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس.

المفاهيم الأساسية للبحث

صدق وثبات الأدوات:

يشير قاسم الصراف(٢٠٠٢، ١٨٩) إلى أن الأدوات لكي تكون مقاييس جيدة يجب أن تتصف بمجموعة من الخصائص من أهمها الصدق والثبات اللذان يعتبران من أكثر المصطلحات الفنية في القياس.

أولاً- صدق الأدوات: تتضمن كتب مناهج البحث عدداً من طرق إيجاد الصدق، منها:

- صدق المحتوى:

يمكن حساب صدق المحتوى للاختبار عن طريق التحليل المبدئي لفقراته بواسطة عدد من المحكمين لتحديد ما إذا كانت هذه الفقرات تتعلق بالجانب المقيس، ثم يقوم الباحث بعمل تكرارات لاستجابات المحكمين ويختار المفردات التي اتفق عليها أكبرعدد من المحكمين. وتشير أدبيات البحث إلى أنه من الأفضل حساب معامل الاتفاق بين المحكمين من خلال معامل اتفاق كندال، ويكون من المناسب استخدام صدق المحتوى مع الاختبارات والاستبيانات التي تقيس مفرداتها التحصيل أو المفاهيم (على عبد الرزاق، ١٩٨٩، ٢٩١؛ على عبد الرزاق وآخرون، ١٩٩٨، ٣٢٨؛ سامية محمد، ٢٠٠٠، ٤٣٩؛ على ماهر، ٢٠٠٣، ١٨٥؛ Gholamreza and Fatemeh, 2008, 152؛ صلاح علام، ٢٠١١، ١٩٢).

- صدق المضمون:

ويطلق عليه البعض صدق الاتساق الداخلي ويمكن حساب صدق المضمون (صدق الاتساق الداخلي) للاختبار بحساب معاملات الارتباط بين درجات الأفراد في كل سؤال ودرجاتهم في الاختبار ككل بعد حذف درجة المفردة من المجموع الكلي للاختبار،وتشير أدبيات البحث إلى أنه من المناسب استخدام صدق المضمون مع الاختبارات واستمارات الملاحظة(سهير بدير، ١٩٨٢،١٣٣؛جون وبست،١٩٨٨،١٩٧؛ حمدي أبو الفتوح،١٩٩٦، ٣٥١ ؛ خير الدين على،١٩٩٧، ٦٣ ؛ محمد السيد، ٢٠٠٠، ٣١٥).

- الصدق التلازمي:

يمكن حساب معامل الصدق بمعامل الارتباط بين درجات الأفراد على الاختبار ودرجاتهم في الأداء الفعلى لجوانب السلوك التي يقيسها الاختبار، وذلك بشرط أن تكون درجات الأداء الفعلية للأفراد قد تم جمعها وقت إجراء الاختبار أو قبلها، ويكون من المناسب استخدام الصدق التلازمي مع الاختبارات التي تقيس جوانب الشخصية بسماتها وأبعادها وصفاتها (محمد عبد السلام، ١٩٩٨، ١٨٨).

- الصدق التنبؤي:

يمكن حساب معامل الصدق في هذه الطريقة بمعامل الارتباط بين الدرجات على الاختبار ودرجات الأداء الفعلي للأفراد كما يقاس بطريقة أخرى بعد إجراء الاختبار بفترة زمنية، وتستخدم هذه الطريقة في حساب صدق اختبارات الاستعدادات الخاصة، مثل: الاستعداد الرياضي أو الاستعداد الميكانيكي، ويطلق على الصدق التلازمي والتنبؤي معاً صدق المحكات، وتشير أدبيات البحث إلى أنه من المناسب استخدام صدق المحكات مع الاختبارات واستمارات الملاحظة والاستبيانات والمقابلات (جون وبست، ١٩٨٨، ١٩٧؛ عزيز حنا وأنور عبد الرحمن ومصطفى كامل، ١٩٩١، ٢١٤؛ حمدي أبو الفتوح، ١٩٩٦، ٣٥١؛ جابر عبد الحميد وأحمد خيري، ١٩٩٦، ٢٤٩؛ أحمد محمد، ٢٠٠٠، ١٢٧؛ بشير صالح، ٢٠٠٠، ١٧٣).

- صدق التكوين الفرضي:

ويطلق عليه أحياناً صدق المفهوم، ويمكن حساب معامل صدق الاختبار بهذه الطريقة بتحديد معامل الارتباط بين درجات الأفراد على الاختبار، وبين مفهوم هذه الجوانب، كما تحددها النظرية التي يتبناها الباحث في أثناء بنائه لهذا الاختبار وتشير أدبيات البحث إلى أنه من المناسب استخدام صدق المفهوم مع المقاييس التي تقيس الاضطرابات الشخصية والسلوكية والحالات المرضية (روبرت ثورندايك وإليزابيث هيجن، ١٩٨٩، ٧١؛ محمود عبد الحليم، ١٩٩٤، ٢٠٨).

- الصدق العاملي:

وتعتمد هذه الطريقة في حساب معامل صدق الاختبار على أسلوب التحليل العاملي، الذي يهدف إلى تحديد مدى قياس مجموعة اختبارات لبعض العوامل المشتركة، ولعلنا نلاحظ البعض يجمع بين الصدق التطابقي والصدق العاملي عند الحديث عن صدق المفهوم، وبالرغم من أهمية طريقة التحليل العاملي إلا أن بعض الباحثين قد يخطئ في استعمال هذه الطريقة (Martin, 2011, 187).

- الصدق الذاتي:

يقاس الصدق الذاتي عن طريق الجذر التربيعي لمعامل ثبات الاختبار (مصطفى باهي وفاتن النمر،٢٠٠٩، ١٠٧)، ومن أهم عيوب معامل الصدق الذاتي أنه يعتمد على معامل الثبات، حيث يشير صفوت فرج(١٩٨٩، ٦٢) إلى أن هذا الأسلوب يتجاهل تماماً المبدأ الأساسي الذي يربط بين مفهومي الصدق والثبات وهذا المبدأ الذي يرى أن كل اختبار صادق ثابت، وليس كل اختبار ثابت صادق، وخلاصة القول أن أي معامل سواء أكان للثبات أم للصدق يجب أن يؤخذ بدقة وبحذر وأن يكون المعامل المستخدم مناسباً لطبيعة البحث، فزيادة معاملات الصدق والثبات ضرورية للبحث في أي نموذج (Nahid, 2003, 602) .

ثانياً– ثبات الأدوات:

الثبات من أهم الشروط السيكومترية للأدوات حيث لا غنى من حسابه مع الصدق؛ لأنه يتعلق بمدى دقة الأدوات في قياس ما تدعي قياسه، ويعد الثبات اتساق أداء الأفراد عبر الزمن إذا ما طبقت عليهم الأداة أكثر من مرة أو هو استقرار أداء الأفراد عبر صور متكافئة من الاختبار، ونستدل أيضا على الثبات من خلال حساب النسبة بين التباين الحقيقي إلى التباين المشاهد (التباين الكلي) لدرجات الاختبار (Kimmo, 2000, 25؛ على ماهر،٢٠٠٣، ١٦٥).

طرق حساب الثبات:

تتعدد أساليب حساب الثبات، ويختص كل أسلوب منها بتقدير نوعية محددة من" تباين الخطأ"، وهو التباين الذي يؤثر على ثبات القياس الذي نحصل عليه كلما استخدمنا مقاييسنا المختلفة، وأشار كل من (محمود عبد الحليم، ١٩٩٤،٢٠٣؛ رجاء علام، ١٩٩٨،٤٢٣؛على ماهر، ٢٠٠٣، ١٦٧؛ بشرى إسماعيل، ٢٠٠٤،٧٢؛ صلاح علام، ٢٠٠٧،٢٣٤) إلى أنه توجد أكثر من طريقة لحساب معامل ثبات الاختبار وهي:

أ– إعادة تطبيق الاختبار. ب– الصور المتكافئة أو البديلة.

جـ– التجزئة النصفية. د– التناسق الداخلي. هـ– تحليل التباين.

الحد المقبول للصدق والثبات:

اعتاد بعض الباحثين على اعتبار الأداة صادقة وثابتة إذا كان معامل الصدق أو معامل الثبات (باعتباره معامل ارتباط) دالاً إحصائياً، إلا أن ذلك لايجوز؛ ويرجع ذلك إلى مفهوم أساسي يرتبط بمدلول معامل الارتباط فعندما يكون معامل الارتباط بين متغيرين ٠,٦ مثلاً، فهذا يعني أن التباين المشترك بين المتغيرين هو ٠,٣٦ (مربع معامل الارتباط)، أي أن نسبة التباين التي يمكن أن نعزوها للمتغير الآخر هي ٣٦%، لنفرض الآن أننا حصلنا على البيانات التالية لاختبار ما، كان معامل الثبات قدره

٠.٠٢، وكان عدد أفراد العينة التي حسب عليها الثبات ١٠٠ فرد، هذه القيمة دالة إحصائياً عند مستوى دلالة ٠.٠١ فهل يمكن القول بأن الاختبار ثابت لأن الارتباط دال إحصائياً عند مستوى ٠.٠١ ؟ بالطبع لا؛ لأن نسبة التباين المشترك التي ترجع إلى ارتباط الاختبار بنفسه لا تزيد عن ٤% وهي نسبة ضئيلة للغاية، وبالتالي لا يمكن اتخاذها قرينة على ثبات الاختبار؛ لذا يجب علينا معرفة الحد المقبول للصدق والثبات، وبالرغم من صعوبة معرفة الحد القبول من صدق الاختبار وثباته، إلا أنه يمكن أن نضع قاعدتين للاسترشاد بهما عند تقويم معاملات صدق الاختبارات وثباتها هما:

١- يجب ألا تقل نسبة التباين المشترك عند حساب معامل ارتباط الاختبار بنفسه (معامل الثبات) عن ٥٠ % وهذا يعني أن الحد الأدنى للثبات يجب أن يكون ٠.٧٠١ (الجذر التربيعي لنسبة التباين المشترك).

٢- يجب أن تكون معاملات الثبات أعلى ما يمكن، فإذا زادت عن ٠.٨٠ كان هذا مفضلاً؛ وذلك عندما يتعلق الأمر بالمقاييس التي تتناول الاتجاهات والميول والنواحي الانفعالية والشخصية، أما إذا كان معامل الثبات يتعلق باختبارات معرفية كاختبارات التحصيل والاستعدادات فهذه يجب أن تكون قريبة جداً من ٠.٩٠ ويفضل أن تزيد عن هذه القيمة، وبالنسبة لمعاملات الصدق فنظراً لأننا نحاول الربط بين الاختبار ومحك لا يمكن اعتباره اختبارا مكافئاً، فإن الحصول على معاملات تزيد عن ٠.٥٠ أو ٠.٦٠ قد يكون أمراً صعباً، ولكن أياً كانت الحالات يجب الحصول على معاملات تعطينا نسبة لا بأس بها من التباين المشترك بين الاختبار والمحك (محمود حسن،١٩٩٦، ١٥٤؛ حسن زيتون، ٢٠٠٨، ١٤٥؛ صلاح علام،٢٠١١، ١٢١).

الدراسات السابقة:

إن المستقرئ للإطار النظري والدراسات السابقة يجد تنوع بين الدراسات العربية والأجنبية المتعلقة بموضوع البحث ومن خلال استعراض هذه الدراسات يمكن ملاحظة ما يلي:

١. تنوعت الدراسات السابقة بين دراسات عربية وأجنبية حيث أكدت دراسة كل من فؤاد البهي (١٩٨٢)، ودراسة عبد العاطي الصياد(١٩٨٨)، ودراسة Michael and Kenneth (2006) على وجود عدة مشكلات في البحث التربوي أدت إلى خلل في البيانات الإحصائية.

٢. توصلت بعض الدراسات إلى الاستخدام غير المناسب للأساليب الإحصائية مثل دراسة أحمد عودة (١٩٩٢)، ودراسة فتحية العجلان (١٩٩٠)، ودراسة عبد الله النجار(١٩٩١)، ودراسة انتصار عبد الرحمن(٢٠٠٨)، ودراسة Goodwin and Goodwin (1985).

٣. كشفت بعض الدراسات مثل دراسة علي الثبيتي(١٩٩٢)، ودراسة رجاء نور(١٩٩٢)، ودراسة أحمد عودة وأحمد الخطيب(١٩٩٤) عن سوء استخدام الأساليب الإحصائية وعدم مناسبتها مع تصميم الدراسة.

٤. أكدت دراسة عبد العزيز عبد الرحمن وشكري أحمد(١٩٩٥) من خلال المقارنة بين الأبحاث العربية والأجنبية على دقة استخدام الأبحاث الأجنبية للأسلوب الإحصائي المناسب مع تقديم التفسير المناسب لذلك.

٥. أشارت بعض الدراسات بضرورة التأكد من شروط استخدام بعض الأساليب الإحصائية كأسلوب التحليل العاملي وتحليل التباين من هذه الدراسات دراسة محمد الشمراني (٢٠٠٠) ودراسة عبد المجيد المالكي(٢٠٠٠).

٦. أكدت دراسة سامية شحاتة(٢٠٠٣) على ضرورة مراعاة بعض المعايير عند استخدام أسلوب صدق المحكمين للتأكد من صدق أدوات البحث، وهذا ما سنؤكد عليه عند تحليلنا لرسائل الماجستير والدكتوراه.

٧. توصلت دراسة كل من صلاح الدين بخيت، وفضل المولى الشيخ(٢٠٠٣)، ودراسة انتصار عبد الرحمن(٢٠٠٨)، ودراسة أشرف السيد(٢٠١١) إلى عدم الاستعمال المناسب للأساليب الإحصائية.

فروض البحث:

من خلال ما تم عرضه من الإطار النظري حول متغيرات البحث ونتائج الدراسات السابقة يمكن عرض فروض البحث فيما يلي:

١- لا توجد فروق ذات دلالة إحصائية بين نسب استعمال الباحثين للأساليب الإحصائية المناسبة وغير المناسبة والمستخدمة عند حساب صدق الأدوات البحثية في الرسائل العلمية (عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس.

٢- لا توجد فروق ذات دلالة إحصائية بين نسب استعمال الباحثين للأساليب الإحصائية المناسبة وغير المناسبة والمستخدمة عند حساب ثبات الأدوات البحثية في الرسائل العلمية (عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس.

إجراءات البحث:

أولاً- عينة البحث : اقتصر هذا البحث على رسائل الماجستير والدكتوراه الممنوحة بالأقسام التربوية المختلفة (علم النفس التربوي، المناهج وطرق التدريس، أصول التربية، تربية الطفل) بكلية التربية جامعة أسيوط خلال الفترة من عام ٢٠٠٤ وحتى عام ٢٠٠٩، وفيما يلي جدول(١) يوضح أعداد رسائل الماجستير والدكتوراه بالأقسام التربوية المختلفة بكلية التربية خلال هذه الفترة.

جدول(١)
أعداد الرسائل بالأقسام المختلفة
بكلية التربية في الفترة من ٢٠٠٤حتى ٢٠٠٩

القسم / نوع الدراسة	ماجستير	دكتوراه
علم النفس التربوي	١٨	١٤
المناهج وطرق التدريس	٣٦	٢٢
تربية الطفل	١٠	١
أصول التربية	٢١	١٥
المجموع	٨٥	٥٢

ثانياً- الخطوات الإجرائية للبحث:

تتلخص خُطوات إجراء البحث في الخُطوات الآتية:

- حصر أعداد رسائل الماجستير والدكتوراه بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط كما هو موضح بجدول (١).
- حصر الأساليب الإحصائية الأكثر شيوعاً في الاستخدام من قبل الباحثين بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط، كما هو موضح بجدول (٢).
- تقييم الأساليب الإحصائية الأكثر استخداماً لحساب كفاءة الأدوات المستخدمة بالرسائل العلمية بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط، والوقوف على مدى صحتها ودقتها من عدمه.

ثالثاً- محكات تقييم الأساليب الإحصائية:

تم تقييم الأساليب الإحصائية المستخدمة في حساب صدق وثبات الأدوات بالرسائل العلمية (عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط في ضوء مستويات القياس ومدى مناسبة البيانات بالرسائل العلمية (عينة الدراسة) بطبيعة هذه المستويات، وتم مراعاة محكين يرتبطين بهذه المستويات هما:

أ- حجم عينة الدراسة.

ب- طبيعة المنهج البحثي المتبع داخل كل قسم تربوي.

جدول(٢)

عدد ونسبة الأساليب الإحصائية المستخدمة

لتحليل بيانات رسائل الماجستير والدكتوراه بالأقسام التربوية بكلية التربية (ن= ١٣٧)*

الأسلوب الإحصائي المستخدم	الصدق		الثبات	
	التكرار	النسبة المئوية	التكرار	النسبة المئوية
اختبار "ت"	٢٦	%٨,١٢	—	———
نسبة الاتفاق بين المحكمين	١٧٩	%٥٦	٣٣	%١٢,٥
معامل ارتباط بيرسون	٥٢	%١٦,٣	١٠٠	%٣٧,٧
معادلة ألفا كرونباك للثبات	—	———	٥٤	%٢٠,١
معامل سبيرمان للرتب	١	%٠,٣١	٥٧	%٢١,٥
التحليل العاملي	٣٠	%٩,٤	—	———
معادلة الصدق الذاتي	٢٥	%٧,٨	—	———
معادلة كيودر – ريتشارد سون	—	———	٢١	%٧,٩
معادلة سكوت للثبات	٥	%١,٥٦	—	———
معادلة هولستي	٢	%٠,٦٢	—	———
مجموع تكرارات الأساليب الإحصائية	٣٢٠	%١٠٠	٢٦٥	%١٠٠

* ن عدد رسائل الماجستير والدكتوراه بالأقسام التربوية الأربعة بكلية التربية والذي يساوي ١٣٧ رسالة.

نتائج البحث وتفسيرها:

أولاً– نتائج الفرض الأول والذي ينص على أنه " لا توجد فروق ذات دلالة إحصائية بين نسب استعمال الباحثين للأساليب الإحصائية المناسبة وغير المناسبة والمستخدمة عند حساب صدق الأدوات البحثية في الرسائل العلمية (عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس".

وللتحقق من صحة هذا الفرض تم تحليل أسلوب الصدق المتبع في أدوات الرسائل العلمية (عينة الدراسة) طبقاً لمستوى قياس البيانات بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط، ثم حساب الفروق بين نسب استعمال الباحثين لهذه الأساليب باستخدام مربع كآي، وقد تبين وجود فروق جوهرية بين مناسبة الاستخدام وعدم مناسبته عند استعمال الأساليب الإحصائية في حساب صدق الأدوات البحثية بالرسائل العلمية (عينة الدراسة) وجدول (٣) يوضح نسبة الأساليب الإحصائية المناسبة وغير المناسبة المستخدمة في تحليل رسائل الماجستير والدكتوراه بالأقسام التربوية المختلفة بكلية التربية جامعة أسيوط.

جدول (٣)

عدد ونسبة الأساليب الإحصائية المناسبة وغير المناسبة المستخدمة في حساب صدق الأدوات برسائل الماجستير والدكتوراه بالأقسام المختلفة بكلية التربية وعددها ن= ١٣٧

القسم التربوي	عدد الرسائل	ع١*	ع٢	نسبة ع١	نسبة ع٢	كا٢	مستوى الدلالة ٠,٠١
تربية الطفل	١١	٦	٢٢	%٢١,٥	%٧٨,٥	١٣,٧	دال
أصول التربية	٣٦	٦	٣٦	%١٤	%٨٦	٢١,٤	دال
المناهج وطرق التدريس	٥٨	٨	١٤٩	% ٥	%٩٥	١٢٦,٦	دال
علم النفس التربوي	٣٢	٥٢	٢٣	%٦٩	%٣١	١١,٢١	دال
المجموع	١٣٧	٧٢	٢٣٠	%٢٤	%٧٦	٨٢,٦٦	دال

* ع١عدد الأساليب الإحصائية المناسبة، ع٢ عدد الأساليب الإحصائية غير المناسبة.

يتضح من جدول (٣) ما يلي:

١- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسب استعمال الباحثين بقسم تربية الطفل للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة صدق الأدوات، وذلك لصالح الاستعمال غير المناسب.

٢- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسب استعمال الباحثين بقسم أصول التربية للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة صدق الأدوات، وذلك لصالح الاستعمال غير المناسب.

٣- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسب استعمال الباحثين بقسم المناهج وطرق التدريس للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة صدق الأدوات، وذلك لصالح الاستعمال غير المناسب.

٤- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسب استعمال الباحثين بقسم علم النفس التربوي للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة صدق الأدوات، وذلك لصالح الاستعمال المناسب.

٥- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسب استعمال الباحثين بالأقسام التربوية المختلفة للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة صدق الأدوات، وذلك لصالح الاستعمال غير المناسب بقسم المناهج وطرق التدريس.

مناقشة نتائج الفرض الأول:

يتضح من جدول (٣) أن نسب استخدام الباحثين للأساليب الإحصائية المناسبة عند حساب صدق الأدوات بأقسام: تربية الطفل، أصول التربية، المناهج وطرق التدريس، علم النفس التربوي كانت على الترتيب: ٢١,٥%، ١٤%، ٥%، ٦٩% بينما كانت نسب استخدام الباحثين للأساليب الإحصائية غير المناسبة عند حساب صدق الأدوات بهذه الأقسام على الترتيب: ٧٨,٥%، ٨٦%، ٩٥%، ٣١% ومن الملاحظ ارتفاع نسبة الأساليب

الإحصائية المناسبة بقسم علم النفس التربوي فهو يتصدر قائمة الأقسام بكلية التربية في دقة ومناسبة استخدامه للأساليب الإحصائية المختلفة؛ حيث بلغت نسبة الأساليب الإحصائية المناسبة عند حساب صدق الأدوات ٦٩ % وهي نسبه مرتفعة، ربما يرجع ذلك إلى طبيعة الدراسة بقسم علم النفس التربوي، واعتماد باحثيه على أبجديات القياس النفسي والتقويم التربوي (جابر عبد الحميد وأحمد خيري، ١٩٩٦، ١٩٤؛ فؤاد البهي، ٢٠١١، ١٢).

وأما عن الأقسام التربوية الأخرى، فكانت نسب الاستخدام غير المناسب للأساليب الإحصائية المتبعة عند حساب صدق الأدوات مرتفعة فقد بلغت هذه النسب بأقسام (تربية الطفل،أصول التربية،المناهج وطرق التدريس)على الترتيب(٧٨,٥%،٨٦%، ٩٥%) وقد يرجع ذلك إلى عدم مراعاة الباحثين بهذه الأقسام لمستويات القياس أو لطبيعة المناهج البحثية المتبعة بكل قسم تربوي، حيث يستخدم قسم أصول التربية المنهج التاريخي، أما قسما تربية الطفل، المناهج وطرق التدريس فيستخدمان المنهج الوصفي التحليلي.

ويمكن تفسير ذلك: بأن الباحثين بالأقسام التربوية مع اختلاف أقسامهم التربوية وبالتالي طبيعة المنهج البحثي المستخدم في الدراسات داخل كل قسم تكون مختلفة، مما يفرض على الباحثين الالتزام ببعض الأساليب الإحصائية التي تكون متوافقة مع طبيعة هذا المنهج، إلا أنه تبين أن بعض الباحثين لا يلتزم بطبيعة المنهج المتبع داخل قسمه التربوي وكان من الأفضل استخدام بعض الأساليب الإحصائية بدلاً من الأساليب المستخدمة في تحليل بيانات الرسائل بالأقسام التربوية الأربعة مع الأخذ في الاعتبار تطبيق شروط كل أسلوب إحصائي، فيُلاحظ أن الباحثين بالأقسام التربوية يستخدمون معامل ارتباط الرتب لسبيرمان مع العلم بأن العينة اعتدالية وحجمها أكبر من ٣٠، فكان من الأنسب استخدام معامل ارتباط بيرسون والذي يطبق في إيجاد الارتباط للعينات البارامترية، ويستخدم كثير من الباحثين أسلوب التحليل العاملي في دراساتهم سواء أكان للتحقق من الصدق العاملي لأدوات الدراسة، أم للكشف عن أقل عدد من العوامل التي تفسر مجموعة من المتغيرات المتعلقة بظاهرة معينة يهتمون بدراستها، غير أنه ربما يساء استخدام هذا الأسلوب؛ وذلك بسبب قلة خبرة بعض الباحثين في المجالات النفسية والتربوية والاجتماعية، حيث يستخدم الباحثون أسلوب التحليل العاملي للتأكد من صدق الأدوات بالرسائل العلمية، مع عدم استيفاء شروطه، وكان من الأفضل استخدام صدق المحكات أو التحقق من شروط أسلوب التحليل العاملي.

ويُلاحظ وجود عدد كبير من الباحثين بقسم المناهج وطرق التدريس يعتمد على الأساليب الإحصائية التي استخدمها من سبقهم من زملائهم الباحثين مع اختلاف طبيعة الأدوات ومستوى قياسها واختلاف حجم العينة؛ مما يؤدي إلى وجود نسبة كبيرة من الأدوات لم يحسب لها الصدق، أيضاً يستخدم الباحثون بهذا القسم معادلة الصدق الذاتي لحساب صدق أدواتهم، والتي تقوم على أستخراج الصدق من معامل الثبات وهذا يتنافى مع طبيعة الشروط السيكومترية للأدوات والتي تؤكد أن كل اختبار صادق ثابت وليس بالضرورة أن يكون العكس صحيحاً، فقد يكون الاختبار ثابت وعند إيجاد صدقه نجده لا يقيس ما وضع لأجله، أضف إلى هذه الأسباب وجود بعض الأدوات لم تطبق لها أية طريقة من طرق حساب الصدق، وهذا خلل في بناء الأدوات ويجب أن يعدل.

ويلاحظ ارتفاع نسبة الاستخدام غير المناسب للأساليب الإحصائية المستعملة في حساب صدق الأدوات بالأقسام التربوية الأربعة، حيث بلغت نسبة الاستخدام غير المناسب ٧٦% وهي نسبة كبيرة مما يستوجب على الباحثين الاهتمام بذلك الأمر فالصدق من أهم الشروط السيكومترية لأية أداة، فهو الأساس في الاعتماد على بناء أية أداة، ولعل السبب في ارتفاع هذه النسبة من الاستعمال غير المناسب أن الباحثين اعتمدوا كثيراً في حساب صدق الأدوات على نسبة الاتفاق بين المحكمين مع العلم بأن هذه النسب يشوبها عدم الدقة؛ وذلك لأنها تُغفل عوامل الصدفة، كما أن الخبراء أو المحكمين في كثير من الأحيان يكونون مشغولون بأبحاثهم ومحاضراتهم وبالتالي عند عرض الباحث الأداة على المحكم لإبداء الرأي في فقرات هذه الأداة، لا يتمكن المحكم أو الخبير من فحص فقراتها نظراً لانشغاله بأبحاثه أو محاضراته؛ لذا كان من الأفضل استخدام صدق المحتوى باستخدام معادلة صدق المحكمين والتي أشار إليها فؤاد أبو حطب وآمال صادق(٢٠٠٨، ١٧٥) وصورتها:

$$C.V.R = \frac{Ne - N/2}{N/2}$$

حيث Ne نسبة الاتفاق بين المحكمين، N/2 نصف عدد المحكمين.

وإذا اعتمدنا على آراء المحكمين بميزان تقدير للمفردات، فيمكن التحقق من صدق المحتوى بتحليل التقديرات إحصائياً باستخدام أساليب تحليل التباين للتحقق من درجة اتساق تقديرات المحكمين، كما يمكن أيضاً التحقق من صدق المحتوى إمبريقياً، وذلك بتطبيق الاختبار على مجموعة من الطلاب قبل بدء عملية التعليم، وإعادة تطبيقه عقب الانتهاء منها، وفحص نتائج الاختبار في المرتين لمعرفة ما إذا كان الاختبار يقيس بالفعل المجال الذي اهتمت به عملية التعليم (صلاح علام ،٢٠١١، ١٩٢).

وأيضاً استخدم بعض الباحثين أسلوب المقارنة الطرفية للتأكد من صدق الأدوات بالرسائل العلمية، مع الوقوع في خطأ شائع ألا وهو حساب دلالة الفروق بين متوسطات الأقوياء والضعفاء على الاختبار، والأصح والأدق أن نقارن بين الإرباعي الأعلى والأدنى على المحك (الميزان) وليس على الاختبار (فؤاد البهي، ٢٠١١، ٤٠٦).

وأيضاً يستخدم بعض الباحثين عدداً كبيراً من الأدوات ظناً منهم أن ذلك يزيد من قيمة البحث وهذا اعتقاد خاطيء كما أكدت على ذلك دراسة (لورنس بسطا وسليمان محمد (١٩٩٤)، ودراسة محمد المُري(٢٠١١) والتي أشارتا إلى أن استخدام الدارسين لعدد كبير من الأدوات لا يزيد من قيمة البحث، بل يكون البحث أكثر عرضة للأخطاء.

ويتضح من خلال التحليل السابق وعرض نسب الأساليب الإحصائية المناسبة وغير المناسبة كما في جدول(٣) وجود فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسب استعمال الباحثين للأساليب الإحصائية المناسبة وغير المناسبة والمستخدمة عند حساب صدق الأدوات البحثية في الرسائل العلمية (عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط في ضوء مستويات القياس.

ومن هذا التحليل نستطيع القول بأن الباحثين لديهم قصور واضح في طرق إيجاد صدق الأدوات حيث بلغت نسبة الاستخدام غير المناسب للأساليب الإحصائية بالأقسام التربوية الأربعة بكلية التربية عند حساب صدق الأدوات ٧٦% كما هو موضح بجدول(٣).

ثانياً- نتائج الفرض الثاني والذي ينص على أنه "لا توجد فروق ذات دلالة إحصائية بين نسب استعمال الباحثين للأساليب الإحصائية المناسبة وغير المناسبة والمستخدمة عند حساب ثبات الأدوات البحثية في الرسائل العلمية(عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط وذلك في ضوء مستويات القياس".

وللتحقق من صحة هذا الفرض تم تحليل أسلوب الثبات المتبع في أدوات الرسائل العلمية (عينة الدراسة) طبقاً لمستوى قياس البيانات بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط، ثم حساب الفروق بين نسب استعمال الباحثين لهذه الأساليب باستخدام مربع كاي وقد تبين من استخدام الباحثين للأساليب الإحصائية المناسبة وغير المناسبة في الأقسام المختلفة بكلية التربية، وجود فروق جوهرية بين مناسبة الاستخدام من عدم مناسبته لحساب ثبات الأدوات وجدول (٤) يوضح نسب الأساليب الإحصائية المناسبة وغير المناسبة المستخدمة في تحليل رسائل الماجستير والدكتوراه بالأقسام المختلفة بكلية التربية.

جدول (٤)

عدد ونسبة الأساليب الإحصائية المناسبة وغير المناسبة المستخدمة في حساب ثبات الأدوات برسائل الماجستير والدكتوراه بالأقسام المختلفة بكلية التربية وعددها ن= ١٣٧

القسم	عدد الرسائل	ع١*	ع٢	نسبة ع١	نسبة ع٢	كا٢	مستوى الدلالة
تربية الطفل	١١	٨	٢٠	٢٩%	٧١%	٥,١٤	دال عند ٠,٠٥
أصول التربية	٣٦	٣٣	٧	٨٢,٥%	١٧,٥%	١٦,٩	دال عند ٠,٠١
المناهج وطرق التدريس	٥٨	٧٤	٥٧	٥٦%	٤٤%	٢,٢	غير دال
علم النفس التربوي	٣٢	٤٣	٣٦	٥٤%	٤٦%	٠,٦٢	غير دال
المجموع	١٣٧	١٥٨	١٢٠	٥٧%	٤٣%	٥,١٩	دال عند ٠,٠٥

* ع ١ عدد الأساليب الإحصائية المناسبة، ع٢ عدد الأساليب الإحصائية غير المناسبة.

يتضح من جدول(٤) ما يلي:

١- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠٥ بين نسبة استعمال الباحثين بقسم تربية الطفل للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة ثبات الأدوات، لصالح الاستعمال غير المناسب.

٢- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسبة استعمال الباحثين بقسم أصول التربية للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة ثبات الأدوات، لصالح الاستعمال المناسب.

٣- لا توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسبة استعمال الباحثين بقسم المناهج وطرق التدريس للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة ثبات الأدوات.

٤- لا توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠١ بين نسبة استعمال الباحثين بقسم علم النفس التربوي للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة ثبات الأدوات.

٥- توجد فروق ذات دلالة إحصائية عند مستوى ٠,٠٥ بين نسبة استعمال الباحثين بالأقسام التربوية المختلفة للأساليب الإحصائية المناسبة وغير المناسبة عند معالجة ثبات الأدوات، لصالح الاستعمال المناسب بقسم أصول التربية.

مناقشة نتائج الفرض الثاني:

يُلاحظ من جدول (٤) أن نسب استخدام الباحثين للأساليب الإحصائية المناسبة بأقسام: تربية الطفل، أصول التربية، المناهج وطرق التدريس، علم النفس التربوي كانت على الترتيب: ٢٩%، ٨٢,٥ %، ٥٦%، ٥٤%، بينما كانت نسب استخدام الباحثين للأساليب الإحصائية غير المناسبة عند حساب ثبات الأدوات بهذه الأقسام على الترتيب: ٧١%، ١٧,٥%، ٤٤%، ٤٦% ويتضح من هذه النسب ارتفاع نسبة الاستخدام المناسب للأساليب الإحصائية بقسم أصول التربية فهو يتصدر قائمة الأقسام بكلية التربية في دقة ومناسبة استخدامه للأساليب الإحصائية المختلفة؛ حيث بلغت نسبة الأساليب الإحصائية المناسبة عند حساب ثبات الأدوات ٨٢,٥ % وهي نسبه مرتفعة، وقد يرجع السبب وراء

ذلك إلى أن أغلب الباحثين بهذا القسم يستخدمون معادلة الاحتمال المنوالي عند حساب ثبات أدواتهم والتي تتناسب وطبيعة المنهج البحثي التاريخي ومستوى القياس التصنيفي (فؤاد البهي، ٢٠١١، ٤٦٤). وبالرغم من ارتفاع هذه النسبة إلا أن أغلب الباحثين يستخدمون نفس المعادلة، معادلة الاحتمال المنوالي لحساب ثبات الأدوات والتي تعتمد في جوهرها على حساب ثبات أسئلة الاستفتاءات التي تعتمد إجابات مفرداتها على اختيار إجابة واحدة صحيحة من عدة إجابات، ولكن هناك بعض الاستبيانات تكون متعددة الاختيارات ولا يوجد بها إجابه صحيحة وأخرى خاطئة كأن تكون بدائل الاستبيان (موافق، إلى حدما، غير موافق) وفي هذه الحالة يكون من الأفضل استخدام معادلات تحليل التباين مثل: معادلات كيودر- ريتشاردسون والتي تستخدم في حالة وجود بديلين، أو معادلة ألفا كرونباخ والتي تستخدم عند حساب الأسئلة التي يكون لها أكثر من بديلين (صلاح علام، ٢٠١١، ١٦٥). وعلى العكس من ارتفاع نسبة الاستخدام المناسب بقسم أصول التربية لوحظ انخفاض هذه النسبة بقسم تربية الطفل حيث بلغت نسبة الاستعمال غير المناسب ٧١% وهي نسبة مرتفعة، ويُفسر وجود هذه النسبة بأن بعض الأدوات لم يحسب لها الثبات، والبعض الآخر طبقت عليه أساليب إحصائية غير مناسبة للمنهج الوصفي المتبع بقسم تربية الطفل، ومن هذه الأساليب معامل ارتباط الرتب لسبيرمان المستخدم في طريقة التجزئة النصفية فقد استخدمه بعض الباحثين مع العلم بكبر حجم العينة واعتدالية التوزيع، وكان من الأفضل استخدام طريقة إعادة التطبيق من خلال حساب معامل ارتباط بيرسون، وقد استخدم بعض الباحثين طريقة إعادة التطبيق بمعامل ارتباط بيرسون لحساب ثبات اختبارات الذاكرة وهذا خلل في شروط الطريقة لأن هذه الطريقة كما هو معلوم لا تصلح لاختبارات الذاكرة وكان من الأنسب تطبيق طريقة أخرى كطريقة الصور المتكافئة (صلاح علام ،٢٠١١، ١٤٦). أيضاً استخدم بعض الباحثين - لحساب ثبات أدواتهم - معادلة ألفا كرونباك لتحليل التباين مع العلم بأن هذه الأدوات لها بديلان اثنان فقط وكان من الأفضل استعمال إحدى معادلات كيودر- ريتشاردسون، وأيضاً استخدم بعض الباحثين معادلة كوبر لحساب نسبة الاتفاق بين الملاحظين وكان من الأفضل استخدام معادلة كوهين والتي تم الإشارة إليها عند مناقشة نتائج الفرض الأول.

وأما بالنسبة لقسمي: المناهج وطرق التدريس، علم النفس التربوي فكانت نسب الاستخدام المناسب وغير المناسب قريبة من بعضها مما أدى إلى عدم وجود فروق جوهرية في مناسبة الاستعمال من عدمه، ويمكن تفسير ذلك بأن الباحثين بهذين القسمين اجتهدوا في أن يستخدموا أساليب إحصائية ملائمة للمناهج البحثية المتبعة داخل كل قسم تربوي، وبالرغم من عدم وجود فروق ذات دلالة؛ إلا أن كثيراً من الباحثين بقسم المناهج وطرق التدريس طبق إحدى معادلات تحليل التباين وهي معادلة كيودر- ريتشاردسون لحساب ثبات اختبار يتكون من ثلاثة بدائل أو أكثر وكان من الأدق تطبيق معادلة ألفا كرونباك (صلاح علام، ٢٠١١، ١٦٥). واستخدم أيضاً بعض الباحثين طريقة إعادة التطبيق مع اختبارات الذكاء والذاكرة وبعض الاختبارات التحصيلية وكان من الأفضل والأنسب تطبيق طريقة الصور المتكافئة أو معادلات التجزئة النصفية (صلاح علام، ٢٠١١، ١٤٦). وأيضاً استخدم بعض الباحثين طريقة إعادة التطبيق على عينة مكونة من راشدين بفاصل زمني أقل من أسبوعين وهذا خلل في شروط طريقة إعادة التطبيق.

ويتضح من خلال التحليل السابق وعرض نسب الأساليب الإحصائية المناسبة وغير المناسبة كما في جدول (٤) وجود فروق ذات دلالة إحصائية عند مستوى ٠,٠٥ بين نسب استعمال الباحثين للأساليب الإحصائية المناسبة وغير المناسبة والمستخدمة عند حساب ثبات الأدوات البحثية في الرسائل العلمية (عينة الدراسة) بالأقسام التربوية الأربعة بكلية التربية جامعة أسيوط في ضوء مستويات القياس.

ومن هذا التحليل نستطيع القول بأن الباحثين بقسم تربية الطفل لديهم قصور واضح في طرق إيجاد ثبات الأدوات حيث بلغت نسب الاستخدام غير المناسب للأساليب الإحصائية بهذا القسم ٧١%، في حين تقاربت نسب الاستخدام المناسب وغير المناسب للأساليب الإحصائية بقسمي المناهج وطرق التدريس، علم النفس التربوي كما هو موضح بجدول (٤٢)، أضف إلى ذلك ارتفاع نسبة الاستخدام المناسب للأساليب الإحصائية عند حساب ثبات الأدوات بقسم أصول التربية حيث بلغت نسبة الاستخدام المناسب ٨٢,٥ %، أيضاً بلغت نسبة الاستخدام غير المناسب للأساليب الإحصائية بالأقسام التربوية الأربعة بكلية التربية عند حساب ثبات الأدوات ٤٣% كما هو موضح بجدول(٤).

قائمة المراجع

أولا : المراجع العربية :

١- بشرى إسماعيل(٢٠٠٤). المرجع في القياس النفسي. القاهرة: الأنجلو المصرية.

٢- بشير صالح الرشيدي(٢٠٠٠). مناهج البحث التربوي " رؤية تطبيقية مبسطة ". الكويت: دار الكتاب الحديث.

٣- جابر عبد الحميد جابر، وأحمد خيري كاظم(١٩٩٦). مناهج البحث في التربية وعلم النفس. القاهرة: دار النهضه العربية.

٤- جون وبست(١٩٨٨). مناهج البحث التربوي. ترجمة: عبد العزيز غانم، الكويت: مؤسسة الكويت للتقدم العلمي.

٥- حسن حسين زيتون(٢٠٠٨). أصول التقويم والقياس التربوي "المفهومات والتطبيقات". الرياض: الدار الصولتية للتربية.

٦- حمدي أبو الفتوح عطيفة(١٩٩٦). منهجية البحث العلمي وتطبيقاتها في الدراسات التربوية والنفسية. القاهرة: دارالنشر للجامعات.

٧- خير الدين علي أحمد (١٩٩٧). دليل البحث العلمي. القاهرة: دار الفكر العربي.

٨- رجاء محمود علام (١٩٩٨). مناهج البحث في العلوم النفسية والتربوية. القاهرة: دار النشر للجامعات.

٩- روبرت ثورندايك، إليزابيث هيجن(١٩٨٩). القياس والتقويم في علم النفس والتربية. ترجمة: عبد الله زيد الكيلاني و عبد الرحمن عدس، الأردن: مركز الكتاب الأردني.

١٠-سامية محمد جابر(٢٠٠٠). منهجيات البحث الإجتماعي والإعلامي. الإسكندرية: دار المعرفة الجامعية.

١١-سهير بدير(١٩٨٢). البحث العلمي " تعريفه، خطواته، مناهجه، أدواته، المفاهيم الإحصائية، كتابة التقرير". مصر: دار المعارف.

١٢-صفوت فرج(١٩٨٩). القياس النفسي. ط٢، القاهرة: الأنجلو المصرية.

١٣-صلاح الدين محمود علام (٢٠٠٧). القياس والتقويم التربوي في العملية التدريسية. عمان: دار المسيرة.

١٤-صلاح الدين محمود علام(٢٠١١). القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.

١٥-عزيز حنا داوُد، وأنور حسين عبد الرحمن، ومصطفى محمد كامل(١٩٩١). مناهج البحث في العلوم السلوكية. القاهرة: الأنجلو المصرية.

١٦-علي عبد الرزاق جلبي(١٩٨٩). تصميم البحث الإجتماعي" الأسس والإستراتيجيات ". الإسكندرية: دار المعرفة الجامعية، ص٢٩١.

١٧-علي عبد الرزاق جلبي، ومحمد عاطف غيث، ومحمد أحمد بيومي، وسامية محمد جابر(١٩٩٨). البحث العلمي الإجتماعي " تصميم خطة وتنفيذها ". الإسكندرية: دار المعرفة الجامعية.

١٨-علي ماهر خطاب(٢٠٠٣). علم النفس الفارق. ط٣، القاهرة: الأنجلو المصرية.

١٩-عماد أحمد حسن(٢٠١٠). القياس النفسي والتقويم التربوي للمعلمين بين النظرية والتطبيق. القاهرة: دار السحاب.

٢٠-فؤاد البهي السيد(٢٠١١). علم النفس الأحصائي وقياس العقل البشري. القاهرة: دار الفكر العربي.

٢١-قاسم علي الصراف(٢٠٠٢). القياس والتقويم في التربية والتعليم. القاهرة: دار الكتاب الحديث.

٢٢-محمد السيد علي(٢٠٠٠). علم المناهج" الأسس والتنظيمات في ضوء الموديولات". القاهرة: دار الفكر العربي.

٢٣-محمد عبد السلام أحمد(١٩٩٨) . القياس النفسي والتربوي. القاهرة: النهضه المصرية.

٢٤-محمود عبد الحليم منسي(١٩٩٤). القياس والإحصاء النفسي والتربوي. الإسكندرية: دار المعارف.

٢٥-محمود حسن إسماعيل(١٩٩٦). مناهج البحث في إعلام الطفل. القاهرة: دار النشر للجامعات.

٢٦-مصطفى باهي، وفاتن النمر(٢٠٠٩). التقويم في مجال العلوم التربوية والنفسية " مبادىء، ونظريات، وتطبيقات". القاهرة: الأنجلو المصرية.

ثانياً : المراجع الأجنبية :

1-Gholamreza, J. & Fatemeh, S. (2008).Validity, Reliability and Difficulty indices for Instructor-Built Exam Questiton. Journal of applied quantitative Methhods 3(2), 151-155.

2-Kimmo, V. (2000). Realibilty of Measurement Scales, Tarkkonen's general method superseds Cronbach's alpha, Department of Statistics, Finland: University of Helsinki.

3-Martin, A. (2011). Not as the crew flies: "Styles" of educational measurement in the reception of inferential statistics at Iowa and Minnesota._University of Chicago: Science History Publications Ltd.

4-Michael, L.& Kenneth, P. (2006). Model selection for the rate problem: A comparison of significance testing, Bayesian, and minimum description length statistical inference. University of Adelaide: Department of Psychology, SA 5005, Available online 27 January 2006.

5-Nahid, G. (2003). Understanding Reliability and Validity in Qualitative Research. Canada: University of Toronto, Toronto, Ontario.